



Working Papers

www.cesifo.org/wp

Elections and Market Provision of Information

Christian Bruns

CESIFO WORKING PAPER NO. 4091

CATEGORY 2: PUBLIC CHOICE

JANUARY 2013

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

CESifo

Center for Economic Studies & Ifo Institute

Elections and Market Provision of Information

Abstract

Economists usually think that rational voters have little incentives to acquire costly information. We present a theoretical model to show that, in contrast to this widely held belief, rational voters acquire considerable amounts of information if media technology is available because then they do not condition their informational decisions on being pivotal. The model also shows that the quality of media coverage is inefficiently low because voters have incentives to free-ride. Further, we show how the quality of information depends on the size of the electorate, the prior knowledge of voters and on the technology to produce information.

JEL-Code: D720, D800, H410.

Keywords: elections, information, media.

*Christian Bruns
University of Goettingen
Platz der Goettinger Sieben 3
Germany – 37073 Goettingen
cbruns@gwdg.de*

January 10, 2013

I would like to thank Robert Schwager for very helpful discussions and advice. Many thanks go also to Emre Aytimur, Lars Ehlers, Steffen Huck, Cesar Martinelli, Panu Poutvaara, Andreas Wagener as well as seminar participants at the Workshop on Economics in Tübingen, the 2nd World Congress of the Public Choice Societies in Miami, the WZB, the IIPF Congress in Dresden and the 6th CESifo Workshop on Political Economy for helpful comments and suggestions. This paper is a substantially revised version of an earlier paper whose title was Too Little News – Collective Decision-Making and Market Provision of Information.

“Today marks a significant transition for The New York Times as we introduce digital subscriptions. It’s an important step that we hope you will see as an investment in The Times, one that will strengthen our ability to provide high-quality journalism to readers around the world and on any platform.”¹

1 Introduction

Ever since Anthony Downs’ (1957) seminal book *An Economic Theory of Democracy* it is a widely held belief among economists that rational voters would have little incentives to acquire costly information that would help them make a well-informed voting decision. According to Downs rational voters are not be willing to pay for information because the probability that their vote changes the result of an election is usually negligible. Thus rational voters decide to remain ignorant.

We show, however, that this result of rational ignorance does not hold in a world where mass media technology is available and where voters have a common motive to find out the truth about the qualities of politicians or policies. In contrast to Downs’ rational ignorance hypothesis, voters can acquire a considerable amount of information in such a scenario.

This scenario appears to be realistic: It can happen, for example, that an election is dominated by a single issue such as an economic crisis, fighting crime or providing (local) public goods like roads and schools. Then, voters might share a common desire to select the most competent politician to cope with the task at hand and might want to acquire additional information about available candidates in order to avoid electing a less competent candidate.

Further, the assumption that voters acquire this information from mass media such as television, radio or (online) newspapers appears to be natural. For a single voter it might be a complicated and overly expensive task to gather information himself but nowadays technology allows to run media companies that hire specialized journalists who gather information and then sell this information to a large number of media consumers.

And indeed, the result that media can provide voters with considerable amounts of information seems to fit in well with our everyday life experience where we can observe informative media coverage of politics and with the findings of surveys where voters say that they receive political information from the mass media (see, e.g., Horrigan

¹Arthur Ochs Sulzberger Jr. (Publisher of The New York Times), “A Letter to Our Readers About Digital Subscriptions”, The New York Times, March 17, 2011, <http://www.nytimes.com/2011/03/18/opinion/18times.html>.

et al. 2004). These observations could not be explained in the case of ignorant voters because commercial mass media then would have no incentive to provide information.

We develop our results in a theoretical model of a two-alternative election. Voters have identical preferences and want to select the best politician but they are uncertain about the talent of the given candidates. As a novel feature, the model explains the production and acquisition of additional information about the candidates instead of assuming exogenous information structures: There are media firms (henceforth: newspapers) that gather information and offer it to the voters. Thus, a voter can either stick to his original information or subscribe to a newspaper in order to acquire more information.

It follows that there can be two groups of voters: subscribers who are relatively well informed and non-subscribers who are less informed. It is shown that non-subscribers abstain in the election so that the votes of the subscribers decide the election and all voters benefit from the well-informed decision of the subscribers. This behavior of non-subscribers is a version of the famous *swing voter's curse* introduced by Feddersen and Pesendorfer (1996) which has recently also been observed in an experimental study by Battaglini et al. (2010). The swing voter's curse ensures that the electoral decision is based upon the information provided by the newspapers in contrast to a situation where costly information can be wasted because the votes of informed voters are not decisive.

The swing voter's curse has important implications for the characteristics of the information service provided by the newspapers. Assuming that profit-maximizing newspapers provide information only if voters pay a positive price it follows:

- (1) Each piece of information some voter has paid for is used to improve the electoral decision and each voter benefits from this improved decision.
- (2) No voter is willing to buy a costly subscription in order to receive a given message that contains the information other voters have paid for. A voter is willing to buy a costly subscription only if this increases the quality of information contained in the message.

Thus, on the one hand, a newspaper subscription serves to receive a message but, on the other hand, it also constitutes a vehicle to contribute to the quality of the information contained in the message. It also follows that each voter considers the contributions of the other voters when he contemplates whether or not to subscribe and we show that the production and acquisition of information can be interpreted as a game of private provision of the public good *quality of political information* where the newspapers determine the characteristics of available contributions.

We find that newspapers provide voters with positive amounts of information so

that electoral decisions are based on more information than the rational ignorance hypothesis suggests. However, the quality of information still is found to be too low as compared to the social optimum because voters have incentives to free-ride.

Further, we find that the quality of the information service does not depend on the size of the electorate: the larger the electorate, the smaller is the increase of information quality per subscription, but the overall quality of the information service remains constant. The gap between the quality provided and the efficient quality, however, increases in the size of the electorate.

The quality of the information service depends on the production technology of the newspapers and on the prior information of the voters. If gathering information is cheap and/or a given investment in investigations yields better information then the newspapers provide a higher overall quality of information. We also find that more precise prior information can lead to both a higher or lower quality of information and it can also be shown that more prior information can leave voters worse off.

Information acquisition by voters is a topic that has only recently received increasing attention in the literature. To our knowledge, this is the first paper that introduces media to explain how much information voters acquire in order to improve the quality of electoral decisions. Related are the papers by Martinelli (2006, 2007 and 2011) who provides interesting results on information aggregation in large elections. As in our paper, Martinelli considers two-alternative elections where voters share a common interest to select the best alternative and where voters can acquire costly information. In these papers each voter can acquire a private signal but there is no information service that provides a signal to many voters as in our model. Oliveros (2011a,b) also studies the aggregation of endogenous information in elections and the relationship between information and abstention.

Strömberg (2004) proposes an alternative reason why voters are willing to pay for political information: informed voters are better able to privately adapt to future policies. In his model, voters who do not read newspapers or pay attention to television programs are excluded from the benefits of the information offered by the mass media because information helps to optimize private behavior. Due to this excludability, mass media programs are biased towards large groups. A study by Aldashev (2010) suggests that voters also engage in information acquisition because political knowledge is a valuable asset for communication within their social networks.

The idea that voters cannot be excluded from well-informed electoral decisions also plays a key role in the paper on the swing voter's curse by Feddersen and Pesendorfer (1996). They show that it can be an optimal strategy of uninformed voters to abstain from voting in order to let informed voters decide an election. While the level of information is exogenously given in the model by Feddersen and Pesendorfer, a study

by Feddersen and Sandroni (2006) presents a theory that explains why voters acquire costly information. The authors say that pivot probabilities are too small in large elections to explain why voters obtain costly information and therefore they propose a framework where voters are willing to pay for information out of a sense of civic duty.

There is also a link to the literature that is concerned with information acquisition in committees. The decisions-theoretic basis of our paper is similar to the paper by Li (2001) where he shows that conservative decision rules can serve to help overcome the free-riding problem concerning information acquisition among members of a group. Mukhopadhyaya (2003), Persico (2004), Gerardi and Yariv (2008) and Gershkov and Szentos (2009) also study information acquisition in committees. See also Gerling et al. (2005) for an excellent survey on this topic.

We proceed as follows: Section 2 describes the model. In section 3 we explain how voters value additional information and describe the optimal voting behavior before we analyze the equilibrium of information provision in section 4. Section 5 presents comparative statics results and section 6 discusses the robustness of our results and some possible extensions. Section 7 concludes.

2 Description of the model

There are $N + 1$ citizens, two candidates and two newspapers. We assume that N is even. The candidates are labeled A and B and the newspapers are labeled 1 and 2, respectively. The citizens can participate in an election in which the candidate that receives a majority of the votes cast wins. A tie is broken by tossing a fair coin.

Citizens. The utility of each citizen is θ_A if candidate A wins the election and θ_B if candidate B wins. Both parameters, θ_A and θ_B , are random variables and can be interpreted to represent the talent of the respective candidate, for example.

Let $\theta := \theta_A - \theta_B$ denote the difference between the abilities of candidate A and candidate B . Then, each citizen's loss is either

$$\ell_A(\theta) = \begin{cases} 0 & \text{for } \theta \geq 0 \\ |\theta| & \text{for } \theta < 0 \end{cases}$$

if candidate A is elected or

$$\ell_B(\theta) = \begin{cases} \theta & \text{for } \theta \geq 0 \\ 0 & \text{for } \theta < 0 \end{cases}$$

if candidate B is elected. So a citizen does not suffer a loss if the more talented

candidate wins; otherwise each citizen's loss is equal to $|\theta|$. Each citizen believes that θ is a normal random variable with mean 0 and precision² τ_θ .

Before the citizens decide whether or not to participate in the election and which candidate to vote for they can subscribe to a newspaper in order to get access to additional information about the two candidates. We label the subscription decision with $d \in \{0, 1, 2\}$, where 0 stands for not subscribing and 1 and 2 for subscribing to the respective newspaper. Further, the voting decision is $v \in \{A, B, \phi\}$, where ϕ denotes abstaining.

Newspapers. The two newspapers offer the subscription packages (p_1, x_1) and (p_2, x_2) , where p_1 and p_2 denote the respective prices of a subscription and x_1 and x_2 represent an increase in the quality of the information that the respective newspaper produces for each subscription sold. Let D_1 and D_2 denote the number of subscriptions that the respective newspaper sells. Then a citizen who subscribes to a newspaper receives a signal

$$s = \theta + \varepsilon$$

where $\varepsilon \sim N(0, x \cdot \tau_\varepsilon)$ and $x = D_1 \cdot x_1 + D_2 \cdot x_2$. The precision of the signal increases linearly in x which can be interpreted as a sample size. This total sample size depends on the extent of sampling by the two newspapers: Newspaper 1 produces a signal

$$s_1 = \theta + \varepsilon_1 \quad \text{with} \quad \varepsilon_1 \sim N(0, D_1 x_1 \cdot \tau_\varepsilon)$$

and newspaper 2 produces a signal

$$s_2 = \theta + \varepsilon_2 \quad \text{with} \quad \varepsilon_2 \sim N(0, D_2 x_2 \cdot \tau_\varepsilon),$$

where the two noise terms are independently distributed.

The idea behind the signal s is that subscribers have access to both the information produced by the newspaper they are subscribed to and the information produced by the other newspaper. As we intend to describe newspaper coverage over a certain period of time (e.g., a whole electoral term) and not the quality of a single copy only, this description seems to be appropriate. When, for example, rumors had it that the former German President Christian Wulff had enjoyed a series of favors and freebies that were considered unseemly the news magazine *Der Spiegel* and the tabloid *Bild* were the first media that investigated and reported on this topic. Once their reports were published other news outlets could observe and deliver them to their customers.

²The precision is defined to be the reciprocal of the variance (precision=1/variance).

Thus, in this model, subscribing to one newspaper provides access to both signals produced by the newspapers and $s = (s_1 + s_2)/2$ and x are sufficient statistics for the single signals.

The newspapers non-cooperatively announce their offers. They choose (p_1, x_1) and (p_2, x_2) in order to maximize their profits

$$\Pi_1[(p_1, x_1), (p_2, x_2)] = D_1[(p_1, x_1), (p_2, x_2)] \cdot (p_1 - \gamma \cdot x_1)$$

and

$$\Pi_2[(p_1, x_1), (p_2, x_2)] = D_2[(p_1, x_1), (p_2, x_2)] \cdot (p_2 - \gamma \cdot x_2),$$

where γ is the cost of producing one unit of x .

Timing. First, the newspapers announce their subscription packages and then the citizens choose whether or not to subscribe. Afterwards the newspapers produce their signals and subscribers observe the signal s . Finally, each citizen decides to vote for one of the candidates or to abstain.

We analyze symmetric equilibria of the game.

3 Strategic voting and the value of information

When each citizen contemplates his voting decision there can be two groups of citizens: *subscribers* and *non-subscribers*. Subscribers receive a signal s of quality x and update their prior belief whereas non-subscribers only have their prior information. Given our assumptions as to the citizens' prior belief about θ and the technology of the information service the optimal posterior decision is described by the solution to a well-known statistical decision problem (see, e.g., DeGroot 1970 or Pratt et al. 1995):

Lemma 1. *For subscribers, the optimal decision posterior to receiving the signal is*

$$v^*(s) = \begin{cases} A & \text{for } s > 0 \\ B & \text{for } s < 0. \end{cases}$$

Proof. See appendix. □

Further, we can observe two important characteristics of the expected loss subscribers attribute to the optimal posterior decision:

Lemma 2.

- (a) *The expected loss attributed to the optimal posterior decision is always smaller than the expected loss attributed to the optimal prior decision:*

$$E[\ell_{v^*}(\theta|s, x)] < E[\ell_{v^*}(\theta)] \quad \text{for all } s \text{ and } x > 0.$$

- (b) *The expected loss attributed to the optimal posterior decision is strictly decreasing in the quality of the information service:*

$$\frac{\partial}{\partial x} E[\ell_{v^*}(\theta|s, x)] < 0.$$

Proof. See appendix. □

Intuitively, Lemma 2 says that an information service of higher quality allows less risky decisions.

Consider a situation where there is a number of i subscribers and where the group of subscribers is split into $(i - k)$ readers of newspaper 1 and k readers of newspaper 2. This implies that subscribers receive a signal s of quality $x = (i - k) \cdot x_1 + k \cdot x_2$. We assume that subscribers know how many copies the newspapers sell.³ After having received the signal a subscriber attributes an expected loss of $r^*[s, x(i, k)]$ to his preferred candidate and an expected loss of $\bar{r}[s, x(i, k)]$ to the other candidate where

$$r^*[s, x(i, k)] := \begin{cases} E[\ell_A(\theta|s, x(i, k))] & \text{for } s > 0 \\ E[\ell_B(\theta|s, x(i, k))] & \text{for } s < 0 \end{cases}$$

$$\bar{r}[s, x(i, k)] := \begin{cases} E[\ell_B(\theta|s, x(i, k))] & \text{for } s > 0 \\ E[\ell_A(\theta|s, x(i, k))] & \text{for } s < 0. \end{cases}$$

Obviously, we have $r^*[s, x(i, k)] < \bar{r}[s, x(i, k)]$ for $s \neq 0$.

Each subscriber knows that no other citizen is better informed than himself. It follows that, from the perspective of an informed citizen, r^* is the minimal expected loss that can be realized and thus each informed citizen wants that his preferred candidate wins the election. The vote of a single subscriber is relevant for the outcome of the election only if it is pivotal and there are three possible events where the vote of a subscriber is pivotal: (1) his preferred candidate leads by one vote, (2) the unwanted candidate leads by one vote or (3) there is a tie.

³In many newspapers you can find information about the number of copies printed. The assumption is, however, not important for our results but it helps to keep things simple. Assuming that subscribers do not know the circulation would make the proof more cumbersome without adding insight.

Assume that $s > 0$ so that subscribers prefer candidate A . Let π_A and π_B denote the probabilities of the events that the respective candidate leads by one vote and let π_t denote the probability of a tie. Then, for each subscriber the difference in expected loss between voting for A and B is

$$E_{\ell_A} - E_{\ell_B} = \pi_A[r^* - 1/2(r^* + \bar{r})] + \pi_B[1/2(r^* + \bar{r}) - \bar{r}] + \pi_t(r^* - \bar{r}) < 0$$

and the respective difference between voting for A and abstaining is

$$E_{\ell_A} - E_{\ell_\phi} = \pi_B[1/2(r^* + \bar{r}) - \bar{r}] + \pi_t[r^* - 1/2(r^* + \bar{r})] < 0.$$

For $s < 0$ we have the corresponding result that $E_{\ell_B} - E_{\ell_A} < 0$ and $E_{\ell_B} - E_{\ell_\phi} < 0$. Thus it is a strictly dominant strategy of each subscriber to vote according to the signal s . The intuition of this result is straightforward: Subscribers know which candidate should be elected and by following their signal they can increase the probability that the preferred candidate wins.

Non-subscribers are indifferent between the two candidates. They know, however, that subscribers can prefer one candidate and that they can benefit from the knowledge of the subscribers if the latter decide the election. But as non-subscribers do not observe the signal they do not know which candidate is preferred by subscribers. If a non-subscriber voted randomly for one of the candidates he would risk to vote against a better informed subscriber which would decrease the probability that the candidate favored by subscribers wins. We show in the appendix that it is a strictly dominant strategy of non-subscribers to abstain because this ensures that the votes of the subscribers decide the election.

Putting the pieces together, we can describe the voting behavior of the citizens as follows:

Proposition 1. *Subscribers vote according to their signal and non-subscribers abstain:*

$$v_{d=1,2}^* = \begin{cases} A & \text{for } s > 0 \\ B & \text{for } s < 0 \end{cases}, \quad v_{d=0}^* = \phi.$$

Proof. See above and the appendix. □

This result is a simple version of the so-called *swing voter's curse* introduced by Feddersen and Pesendorfer (1996) who provide an information-based rationale for voters to abstain in elections. In our model, the swing voter's curse has important implications for how citizens evaluate the information produced by the newspapers:

It ensures that the electoral decision is based upon all information that has been produced. This implies that each citizen has an expected loss of $r^*(s, x(i, k))$ given a signal characterized by $[s, x(i, k)]$.

The citizens, however, do not know $r^*(s, x(i, k))$ when they consider whether or not to subscribe because neither s nor x have been determined yet. From the prior perspective, the expectation of the posterior expected loss given that there are $i - k$ and k readers of the respective newspapers is

$$E[r^*(s, x(i, k))] = \int_{-\infty}^{\infty} r^*(s, x(i, k)) f(s, x(i, k)) ds.$$

and the value of an information service of quality $x(i, k)$ is the difference between the expected loss without any newspaper coverage (i.e., the electoral decision is based upon the prior belief only) and the prior expectation of the expected loss after having observed the signal:

$$\begin{aligned} \omega[x(i, k)] &:= E[l_{v^*}(\theta)] - E[r^*(s, x(i, k))] \\ &= E[l_{v^*}(\theta)] - E\{E[l_{v^*}(\theta) | S, x(i, k)]\}. \end{aligned}$$

Given the assumptions as to the prior belief and to the technology of the information service, the following can be shown:

Lemma 3. *The value a citizen attributes to an information service of quality x from the prior perspective is described by the function*

$$\omega(x) = \left(\frac{\tau_\theta(\tau_\theta + \tau_\varepsilon x)}{\tau_\varepsilon x} \right)^{-1/2} \cdot \varphi(0),$$

where φ is the density function of the standard normal distribution. The function $\omega(x)$ is strictly increasing in x , strictly concave and has infinite slope at the origin.

Proof. See appendix. □

While a realized value s results from a random process the value of x , the quality of the information service, is determined by the decisions of the newspapers and the citizens. In the next section, we determine the equilibrium quality of information.

4 Market equilibrium

Consider a citizen c in order to analyze a symmetric equilibrium. If the citizen decides not to subscribe then he abstains and benefits from the information provided by other

citizens. For given strategies of the newspapers and of the other citizens the expected benefit from not-subscribing is equal to

$$\tilde{\omega}(x_{-c}) = \sum_{i=0}^N \sum_{k=0}^i q(i, k) \cdot \omega[x(i, k)],$$

where $q(i, k)$, the probability that there are $i - k$ readers of newspaper 1 and k readers of newspaper 2, is equal to

$$q(i, k) = \frac{N!}{(i - k)!k!(N - i)!} \sigma_1^{i-k} \sigma_2^k \sigma_0^{N-i}.$$

Thus, citizen c subscribes to a newspaper if the additional value he derives from subscribing is larger than the price of the subscription, i.e.

$$\tilde{\omega}(x_n + x_{-c}) - \tilde{\omega}(x_{-c}) - p_n > 0$$

for at least one newspaper $n=1,2$. For given offers (p_1, x_1) and (p_2, x_2) the citizen subscribes to newspaper 1 if and only if

$$\tilde{\omega}(x_1 + x_{-c}) - \tilde{\omega}(x_2 + x_{-c}) > p_1 - p_2$$

and to newspaper 2 if and only if

$$\tilde{\omega}(x_1 + x_{-c}) - \tilde{\omega}(x_2 + x_{-c}) < p_1 - p_2.$$

Price equilibrium. Assume that the qualities x_1 and x_2 of the subscriptions are given and consider the decision of citizen c for a given pair of offers: At least one newspaper has always an incentive to undercut its rival's price in order to attract the citizen. In response, the undercut newspaper lowers its price until the citizen favors its offer again and this Bertrand-like price competition pushes prices downwards. We have to distinguish between price-equilibria with equal and unequal given qualities:

- *Unequal qualities.* Assume with no loss of generality that $x_1 > x_2$. If newspaper 1 offers a higher quality, it charges

$$p_1^* = \tilde{\omega}(x_1 + x_{-c}) - \tilde{\omega}(x_2 + x_{-c}) + \gamma x_2$$

and earns a positive profit. If newspaper 1 charges a higher price, newspaper 2 can undercut this price and attract the citizen. However, to ensure non-negative profits, newspaper 2 cannot announce a price lower than $p_2 = \gamma \cdot x_2$ and thus it cannot undercut if newspaper 1 charges p_1^* . Consequently, p_1^* is the profit-

maximizing price for newspaper 1 while the price of newspaper 2 is $p_2^* = \gamma \cdot x_2$.

For $x_1 < x_2$ it follows that

$$p_2^* = \tilde{\omega}(x_2 + x_{-c}) - \tilde{\omega}(x_1 + x_{-c}) + \gamma x_1 \quad \text{and} \quad p_1^* = \gamma \cdot x_1.$$

- *Equal qualities.* In case of equal qualities the newspapers can undercut each other's prices as long as the prices are above the cost to produce a subscription of quality x_1 or x_2 . Thus, the only price equilibrium has both newspapers set equal prices $p_1^* = p_2^*$, which are characterized by

$$p_1^* = \gamma \cdot x_1 \quad \text{and} \quad p_2^* = \gamma \cdot x_2.$$

Having analyzed the price setting stage we can now turn to the stage where the newspapers choose the quality increase per subscription.

Quality equilibrium. The profits from selling a subscription to citizen c are

$$\Pi_{1c}(x_1, x_2) = \begin{cases} p_1^*(x_1, x_2) - \gamma \cdot x_1 & \text{for } x_1 > x_2 \\ 0 & \text{for } x_1 \leq x_2 \end{cases}$$

and

$$\Pi_{2c}(x_1, x_2) = \begin{cases} 0 & \text{for } x_1 \geq x_2 \\ p_2^*(x_1, x_2) - \gamma \cdot x_2 & \text{for } x_1 < x_2. \end{cases}$$

It is a weakly dominant strategy for each newspaper to choose the quality that maximizes the profit per subscription for a given strategy of the other newspaper. Due to the symmetry of the newspapers' maximization problem, both newspapers offer the same quality increase per subscription characterized by

$$\tilde{\omega}'(x_n + x_{-c}) = \gamma. \tag{1}$$

This means that the citizen's marginal willingness to pay for higher quality, evaluated at the aggregate amount of contributions equals the cost of increasing quality by one unit. The assumption of identical citizens implies that, in a symmetric equilibrium, the solution to equation (1) is characterized by a situation where each citizen contributes an identical amount to the quality of the information service so that the total quality adds up to $x = (N + 1) \cdot x_n$. Given the properties of the function $\omega(x)$ as described in Lemma 3 there is a unique solution to equation (1) and we can describe the equilibrium of the model as follows:

Proposition 2. *In the unique equilibrium, the newspaper market provides an information service of quality $x^* = x^*(\gamma, \tau_\varepsilon, \tau_\theta) > 0$ and this quality is implicitly defined by*

$$\omega'(x^*) = \gamma. \tag{2}$$

Each citizen contributes the amount $x_n^ = x^*/(N+1)$ via a subscription and pays the price $p_n^* = \gamma \cdot [x^*/(N+1)]$.*

Proposition 2 shows that the newspapers provide the citizens with information that helps achieve a less risky electoral decision. The quality of this information is determined by the efficiency of the technology of information production (parameters γ and τ_ε) and the precision of the prior belief (τ_θ). Before we describe how variations of these parameters affect x^* , we compare the market result with a situation where a social planner employs the information technology of the newspapers in order to examine whether the market provision of information is efficient. A social planner chooses the quality of the information service to maximize

$$(N+1) \cdot \omega(x) - \gamma \cdot x.$$

Thus, the efficient quality of information, \hat{x} , is implicitly defined by

$$(N+1) \cdot \omega'(\hat{x}) = \gamma. \tag{3}$$

Comparing the equations (2) and (3) we can see that x^* is not efficient. Both equations differ in the factor $(N+1) > 1$ only. As $\omega'(x)$ is a strictly decreasing function, the value of \hat{x} must be larger than the value of x^* . Thus, we can say that:

Proposition 3. *The market provides an information service whose quality is inefficiently low.*

This result is not surprising given the analysis above: Subscriptions to the newspapers constitute a vehicle to contribute to the public good *quality of information* and thus each citizen considers the contribution of others when he determines his willingness to pay for information. So media technology helps overcome the problem of low pivot probabilities of a single vote that can prevent voters from acquiring information but we are still in a classic situation of free-riding. The market outcome corresponds to a game of private provision of a public good where the citizens would individually choose a contribution to the quality of information at a cost of γ per unit.⁴ Thus, our

⁴This insight is analogous to that by Besley and Ghatak (2007) who study the economics of corporate social responsibility.

model reveals a close link between the theory of information acquisition by rational voters and the literature on the private provision of public goods.

In the next section, we further characterize the market provision of information by presenting comparative statics results before we discuss some relevant assumptions of our model.

5 Comparative statics analysis

It remains to describe how changes of the parameters in the model alter the equilibrium outcome. We analyze the changes of both x^* , the quality of the information service, and $\rho_t(x^*)$, the total expected loss each citizen attributes to the electoral decision in equilibrium. The total expected loss is equal to the sum of the expected loss from the prior perspective and the price of a subscription:

$$\begin{aligned}\rho_t(x^*) &:= E[r^*(s, x^*)] + \gamma \cdot x^*/(N + 1) \\ &= E[\ell_{v^*}(\theta)] - \omega(x^*) + \gamma \cdot x^*/(N + 1).\end{aligned}$$

The proofs are relegated to the appendix.

- (a) *Size of the electorate (N).* To characterize the comparative statics with respect to N it suffices to note that the value of x^* described by Proposition (2) is independent of the size of the electorate. Consequently, the quality of the information service remains the same whereas the individual contributions $x^*/(N+1)$ decrease in the size of the electorate. It also follows from Proposition (2) that the total total expected loss of a citizen decreases in N .

This result implies further that the gap between the efficient level \hat{x} and the market level x^* is larger, the larger the electorate is. While the strong result of a constant market level in our setting depends on the quasi-linear structure of the voters' preferences, the problem that the gap between the efficient level and the market level of a public good increases in the size of the population is of a general nature as, e.g., Pecorino (1999) shows.

We can also make an interesting comparison to a result by Strömberg (2004) who finds that mass media allocate more resources to large groups. In his model, voters who do not buy a newspaper are excluded from the benefit of information because the value of information originates from a private action. This is different in our model where non-subscribers cannot be excluded from the benefits of information. Thus, there are no increasing returns to scale in this model.

Once we consider an additional dimension of media coverage, such as entertainment or information in the sense of Strömberg characterized by increasing returns to scale, the lack of this effect might make it less attractive to report on politics so that there might be a tendency that political information is crowded out when the number of voters increases. However, reports on politics could also be a mix of entertaining and informative content and the extent of the former could increase in the number of voters whereas the latter remains constant (A journalist could write a new celebrity kind of story about a politician and add some hard information from the archives, for example). Thus, our model should still have explanatory power when we introduce additional dimensions of media coverage.

So far, we have relied on the concavity of $\omega(x)$ to characterize the market equilibrium. Recalling the functional form of $\omega(x)$ as shown in Lemma 3, equation (2) that implicitly defines x^* , the quality of information in equilibrium, reads

$$\frac{1}{2}\varphi(0)\frac{\tau_\theta^2}{\tau_\varepsilon x^2}\left(\frac{\tau_\theta^2}{\tau_\varepsilon x} + \tau_\theta\right)^{-3/2} - \gamma = 0$$

and we use this equation to derive the remaining comparative statics results.

- (b) *Technology of information production* (γ and τ_ε). Recall that $\omega'(x)$ is a strictly decreasing function. Thus, lower values of γ result in an increase of information production, or $\partial x^*/\partial \gamma < 0$.

The effect of a higher precision of an observation (τ_ε) on x^* is ambiguous:

$$\frac{\partial x^*}{\partial \tau_\varepsilon} \begin{cases} > 0 & \text{for } \tau_\theta > 2\tau_\varepsilon x^* \\ = 0 & \text{for } \tau_\theta = 2\tau_\varepsilon x^* \\ < 0 & \text{for } \tau_\theta < 2\tau_\varepsilon x^*. \end{cases}$$

The equilibrium precision of the information service ($\tau_\varepsilon x^*$), however, increases in τ_ε and it is shown in the appendix that lower values of γ as well as larger values of τ_ε lead to a lower total expected loss of each citizen. So a more efficient production technology benefits the citizens.

- (c) *Precision of the prior belief* (τ_θ). The effect of a more precise prior on informa-

tion production is ambiguous:

$$\frac{\partial x^*}{\partial \tau_\theta} \begin{cases} > 0 & \text{for } \tau_\theta < \tau_\varepsilon x^*/2 \\ = 0 & \text{for } \tau_\theta = \tau_\varepsilon x^*/2 \\ < 0 & \text{for } \tau_\theta > \tau_\varepsilon x^*/2. \end{cases}$$

Though one may intuitively expect that a more informed prior implies less information production this only holds true for relatively 'large' values of τ_θ as compared to $\tau_\varepsilon x^*$. Further, a more precise prior can result in an increase of a voter's total risk (ρ_t), so that we conclude that more prior information can lead to comparatively adverse outcomes.

6 Discussion

Our results are derived in a setting where we make somewhat optimistic assumptions about the market environment while we could say that the assumption of purely non-cooperative behavior of the voters is a pessimistic one. The model constitutes a benchmark case that serves to characterize the provision of information by a competitive media market in absence of other adverse effects from the supply side on the level of information in equilibrium. A short discussion of the assumptions is in order.

The assumption of stiff competition between the newspapers serves to achieve a situation where the voters are not harmed because of market power the newspapers might possess. A monopoly newspaper, for example, would have incentives to charge a price higher than the production costs in order to raise profits. From the voters' perspective, there is no such waste associated with information acquisition in case of a fully competitive market.

We are optimistic concerning the problem that the newspapers commit to collecting the amount of information the voters have paid for. Once the newspapers have received their payments they have incentives to provide a signal that is based on fewer observations in order to raise profits. In case of imperfect monitoring, the newspapers can try to build up reputations of being high quality providers which do not cheat. Besley and Ghatak (2007) show that competing firms need to charge prices above marginal costs to make credible promises in such a setting. This result indicates that the media market produces less information when the newspapers cannot credibly commit to sell information at marginal cost.

Further, the informational situation is assumed to be efficient in such a way that every single piece of information that is produced definitely helps to improve the electoral decision. Suppose, for example, that the voters only receive the information

produced by the newspaper they subscribe to but no other information and that there is no communication between voters. Then, the first-order condition in equilibrium would read $q \cdot \omega(x_n^+) = k$, where $0 \leq q \leq 1$ denotes the probability that the signal provided by a newspaper n is pivotal for the electoral decision. This follows because the payoff of contributing to the quality of the respective signal is $\omega(x_n^+)$ with probability q and zero with probability $1 - q$. Thus, for $q < 1$ all signals are less informative in equilibrium. However, it appears to be more realistic that once information is produced it will spread out because of communication between journalists or voters.

Our assumption that the voters act non-cooperatively yields a pessimistic benchmark. The discussion on public good provision and collective action goes back at least to Olson (1965) and the assumption of non-cooperative behavior appears to be appealing, especially in large groups. However, cooperation can sustain even when groups are large as is shown theoretically by Pecorino (1999), for example. With cooperative behavior, the total level of a public good is higher than in our non-cooperation equilibrium.

7 Concluding remarks

This paper offers a new perspective on the motivation of rational voters to acquire costly information. We have shown that information acquisition needs not be driven by the pivot probabilities of single votes as it is often assumed in the literature. Instead, due to the availability of media technology we can understand the production and acquisition of political information as a game of private provision of a public good in a sense that is closer to the traditional analysis of public goods than previous work on information acquisition by voters.

The result that the quality of information provided by a competitive media market is inefficiently low could be interpreted as a rationale for public provision of political information. We do not jump to this conclusion, however, but we think that answering this question is beyond the scope of the analysis: Think, on the one hand, of the problem that public institutions may also work inefficiently as it is widely discussed in the literature. Moreover, it is a delicate question to which extent government should be allowed to control modern mass media that do not only act as neutral information providers as in our analysis but may shape our thinking about the world in a more general way (see Prat and Strömberg (2011) for a discussion). Thus, we confine ourselves to believing that our analysis points out an important problem and that it constitutes an interesting benchmark case and starting point for further research.

One promising direction for further research is to leave behind the assumption of identical voters and to introduce different tastes for public goods provided by

politicians or ideologically motivated preferences when it comes to selecting political candidates. It appears reasonable that these extensions would alter our results in a quantitative dimension but the main results of the paper that (i) pivot probabilities of single votes are not important and that (ii) the free-riding behavior of voters results in an inefficiently low quality of information should prove robust. Thus, we expect to observe variations of the theme presented in this paper but no substantial deviations.

Further, political candidates enter our model as parameters only but politicians or parties usually have strong incentives to manipulate the voters' beliefs about their talents in order to secure political offices. Introducing politicians as players into the game would provide, for example, an interesting link from our topic of information acquisition of voters to the literature on rent-seeking by politicians.

Proofs

Proof of Lemma 1. Let $F(\theta)$ be the distribution function of θ . A citizen prefers candidate A if

$$E[\ell_A(\theta)] = \int_{-\infty}^0 |\theta| dF(\theta) < E[\ell_B(\theta)] = \int_0^{\infty} \theta dF(\theta).$$

Moving terms on one side shows that this is equivalent to

$$\int_{-\infty}^0 \theta dF(\theta) + \int_0^{\infty} \theta dF(\theta) = E(\theta) > 0.$$

Thus, for $E[\ell_A(\theta)] > E[\ell_B(\theta)] \Leftrightarrow E(\theta) < 0$ citizens prefer candidate B .

Given the assumptions as to the citizens' prior belief and to the technology of the information service, the posterior distribution function of θ is a normal distribution with mean $\mu = \frac{\tau_\varepsilon x}{\tau_\theta + \tau_\varepsilon x} \cdot s$ and precision $\tau = \tau_\theta + \tau_\varepsilon x$ (see, e.g., DeGroot 1970). Considering the mean of the posterior distribution it follows that

$$v^*(s) = \begin{cases} A & \text{for } s > 0 \\ B & \text{for } s < 0. \end{cases}$$

□

Proof of Lemma 2. Posterior expected loss. The posterior expected loss is either

$$E[\ell_A(\theta|s)] = \int_{-\infty}^0 -\theta dF(\theta|s)$$

or

$$E[\ell_B(\theta|s)] = \int_0^{\infty} \theta dF(\theta|s),$$

where $F(\theta|s)$ denotes the posterior distribution function of θ as described above. The integrals on the right-hand side of the equations can be interpreted as *partial expectations* of the random variable θ .

We derive the functional form of such partial expectations in general for a normal random variable W with mean μ and precision τ because we can use this result later. The density function and the distribution function of W are denoted $f(w)$ and $F(w)$, respectively. First, we show that

$$\int_{-\infty}^0 -w f(w) dw = \tau^{-\frac{1}{2}} \left[\varphi(-\tau^{\frac{1}{2}} \mu) - \tau^{\frac{1}{2}} \mu \cdot \Phi(-\tau^{\frac{1}{2}} \mu) \right].$$

We start with substituting $w f(w)$ in the integral on the left-hand side with $\mu f(w) - \frac{f'(w)}{\tau}$ which follows directly from

$$f'(w) = -\tau(w - \mu)f(w).$$

$$\begin{aligned} \int_{-\infty}^0 -wf(w)dw &= - \int_{-\infty}^0 wf(w)dw \\ &= -\mu \int_{-\infty}^0 f(w)dw + \frac{1}{\tau} \int_{-\infty}^0 f'(w)dw \\ &= -\mu \cdot F(0) + \frac{1}{\tau} f(0) \\ &= \tau^{-\frac{1}{2}} \left[\tau^{-\frac{1}{2}} f(0) - \mu \tau^{\frac{1}{2}} \cdot F(0) \right] \end{aligned}$$

Transforming to the standard normal form we obtain

$$\int_{-\infty}^0 -wf(w)dw = \tau^{-\frac{1}{2}} \left\{ \varphi[-\mu\tau^{\frac{1}{2}}] - \mu\tau^{\frac{1}{2}} \cdot \Phi[-\mu\tau^{\frac{1}{2}}] \right\}. \quad (4)$$

Applying an analogous procedure confirms that

$$\int_0^{\infty} wf(w)dw = \tau^{-\frac{1}{2}} \left[\varphi(-\tau^{\frac{1}{2}}\mu) + \tau^{\frac{1}{2}}\mu \cdot (1 - \Phi(-\tau^{\frac{1}{2}}\mu)) \right]. \quad (5)$$

Proof of part (a). We have to prove that

$$\tau_{\theta}^{-\frac{1}{2}} \cdot \varphi(0) > \tau^{-\frac{1}{2}} \cdot \varphi(-\mu\tau^{\frac{1}{2}}) - \underbrace{\mu \cdot \Phi(-\mu\tau^{\frac{1}{2}})}_{\geq 0}$$

and

$$\tau_{\theta}^{-\frac{1}{2}} \cdot \varphi(0) > \tau^{-\frac{1}{2}} \cdot \varphi(-\mu\tau^{\frac{1}{2}}) - \underbrace{|\mu| \cdot [1 - \Phi(-\mu\tau^{\frac{1}{2}})]}_{\geq 0}.$$

It is straightforward to see that $\tau_{\theta} < \tau = \tau_{\theta} + \tau_{\varepsilon}x$. Thus $\tau_{\theta}^{-\frac{1}{2}} < \tau^{-\frac{1}{2}}$. Further, the maximum of a standard normal density function $\varphi(z)$ is at $z = 0$ which means that $\varphi(0) \geq \varphi(-\mu\tau^{\frac{1}{2}})$ and consequently $\tau_{\theta}^{-\frac{1}{2}} \cdot \varphi(0) > \tau^{-\frac{1}{2}} \cdot \varphi(-\mu\tau^{\frac{1}{2}})$. Subtracting a positive term on each right-hand side further reduces the total value of each right-hand side.

Proof of part (b). Straightforward calculations show that

$$\begin{aligned}\frac{\partial E[\ell_A(\theta|s)]}{\partial x} &= \frac{\partial E[\ell_A(\theta|s)]}{\partial \mu} \frac{\partial \mu}{\partial x} + \frac{\partial E[\ell_A(\theta|s)]}{\partial \tau} \frac{\partial \tau}{\partial x} \\ &= -\Phi(-\mu\tau^{\frac{1}{2}}) \cdot \frac{\tau_\theta \tau_\varepsilon s}{(\tau_\theta + \tau_\varepsilon x)^2} - \frac{1}{2} \tau^{-\frac{3}{2}} \varphi(-\mu\tau^{\frac{1}{2}}) \cdot \tau_\varepsilon < 0\end{aligned}$$

and

$$\begin{aligned}\frac{\partial E[\ell_B(\theta|s)]}{\partial x} &= \frac{\partial E[\ell_B(\theta|s)]}{\partial \mu} \frac{\partial \mu}{\partial x} + \frac{\partial E[\ell_B(\theta|s)]}{\partial \tau} \frac{\partial \tau}{\partial x} \\ &= -[1 - \Phi(-\mu\tau^{\frac{1}{2}})] \cdot \frac{\tau_\theta \tau_\varepsilon |s|}{(\tau_\theta + \tau_\varepsilon x)^2} - \frac{1}{2} \tau^{-\frac{3}{2}} \varphi(-\mu\tau^{\frac{1}{2}}) \cdot \tau_\varepsilon < 0\end{aligned}$$

□

Proof of Proposition 1. We define:

$$\begin{aligned}\sigma_1 &:= \text{Prob}[d = 1 \wedge v = A | s > 0] = \text{Prob}[d = 1 \wedge v = B | s < 0], \\ \sigma_2 &:= \text{Prob}[d = 2 \wedge v = A | s > 0] = \text{Prob}[d = 2 \wedge v = B | s < 0], \\ \sigma_A &:= \text{Prob}[d = 0 \wedge v = A], \\ \sigma_B &:= \text{Prob}[d = 0 \wedge v = B], \\ \sigma_\phi &:= \text{Prob}[d = 0 \wedge v = \phi].\end{aligned}$$

Then, the probabilities of the events where the vote of a non-subscriber is pivotal and where there are $i - k$ and k readers of the respective newspapers are

$$\begin{aligned}\pi_A^\oplus(i, k) &= \sum_{j=0}^{N/2-i} \frac{N!}{(i-k)!k!j!(i+j-1)!(N-2i-2j+1)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^j \sigma_B^{i+j-1} \sigma_\phi^{N-2i-2j+1} \\ \pi_A^\ominus(i, k) &= \sum_{j=0}^{N/2-i-1} \frac{N!}{(i-k)!k!j!(i+j+1)!(N-2i-2j-1)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^{i+j+1} \sigma_B^j \sigma_\phi^{N-2i-2j-1} \\ \pi_B^\oplus(i, k) &= \sum_{j=0}^{N/2-i-1} \frac{N!}{(i-k)!k!j!(i+j+1)!(N-2i-2j-1)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^j \sigma_B^{i+j+1} \sigma_\phi^{N-2i-2j-1} \\ \pi_B^\ominus(i, k) &= \sum_{j=0}^{N/2-i} \frac{N!}{(i-k)!k!j!(i+j-1)!(N-2i-2j+1)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^{i+j-1} \sigma_B^j \sigma_\phi^{N-2i-2j+1} \\ \pi_0^\oplus(i, k) &= \sum_{j=0}^{N/2-i} \frac{N!}{(i-k)!k!j!(i+j)!(N-2i-2j)!} \sigma_1^{i-k} \sigma_2^k \sigma_B^i (\sigma_A \sigma_B)^j \sigma_\phi^{N-2i-2j} \\ \pi_0^\ominus(i, k) &= \sum_{j=0}^{N/2-i} \frac{N!}{(i-k)!k!j!(i+j)!(N-2i-2j)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^i (\sigma_A \sigma_B)^j \sigma_\phi^{N-2i-2j}\end{aligned}$$

where \oplus and \ominus denote that $s > 0$ and $s < 0$, respectively. Then, the difference in expected loss between voting for A

and B is

$$\begin{aligned}
Er_A - Er_B &= \sum_{i=1}^{N/2} \sum_{k=0}^i \left\{ \int_0^{\infty} \left[\pi_A^{\oplus}(i, k) \cdot r^*(s, x(i, k)) + \pi_B^{\ominus}(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] + \pi_0^{\oplus}(i, k) \cdot r^*(s, x(i, k)) \right] f(s, x(i, k)) ds \right. \\
&\quad + \int_{-\infty}^0 \left[\pi_A^{\ominus}(i, k) \cdot \bar{r}(s, x(i, k)) + \pi_B^{\oplus}(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] + \pi_0^{\ominus}(i, k) \cdot \bar{r}(s, x(i, k)) \right] f(s, x(i, k)) ds \\
&\quad - \int_0^{\infty} \left[\pi_A^{\oplus}(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] + \pi_B^{\oplus}(i, k) \cdot \bar{r}(s, x(i, k)) + \pi_0^{\oplus}(i, k) \cdot \bar{r}(s, x(i, k)) \right] f(s, x(i, k)) ds \\
&\quad \left. - \int_{-\infty}^0 \left[\pi_A^{\ominus}(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] + \pi_B^{\ominus}(i, k) \cdot r^*(s, x(i, k)) + \pi_0^{\ominus}(i, k) \cdot r^*(s, x(i, k)) \right] f(s, x(i, k)) ds \right\}.
\end{aligned}$$

Due to the symmetry of the prior and the signal it can be shown that $f(s = a, x(i, k)) = f(s = -a, x(i, k))$, $r^*(s = a, x(i, k)) = r^*(s = -a, x(i, k))$ and $\bar{r}(s = a, x(i, k)) = \bar{r}(s = -a, x(i, k))$. Thus we can define

$$r^*(x(i, k)) := \int_0^{\infty} r^*(s, x(i, k)) f(s, x(i, k)) ds = \int_{-\infty}^0 r^*(s, x(i, k)) f(s, x(i, k)) ds$$

and

$$\bar{r}(x(i, k)) := \int_0^{\infty} \bar{r}(s, x(i, k)) f(s, x(i, k)) ds = \int_{-\infty}^0 \bar{r}(s, x(i, k)) f(s, x(i, k)) ds.$$

It follows that

$$Er_A - Er_B = \sum_{i=1}^{N/2} \sum_{k=0}^i \underbrace{[r^*(i, k) - \bar{r}(i, k)]}_{<0} \cdot \left\{ \frac{1}{2} \left[\pi_A^{\oplus}(i, k) - \pi_B^{\ominus}(i, k) + \pi_B^{\oplus}(i, k) - \pi_A^{\ominus}(i, k) \right] + \pi_0^{\oplus}(i, k) - \pi_0^{\ominus}(i, k) \right\},$$

where

$$\begin{aligned}
\pi_A^{\oplus}(i, k) - \pi_B^{\ominus}(i, k) &= \sum_{j=0}^{N/2-i} \frac{N!}{(i-k)!k!j!(i+j-1)!(N-2i-2j+1)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^j \sigma_B^j \sigma_{\phi}^{N-2i-2j+1} \cdot (\sigma_B^{i-1} - \sigma_A^{i-1}), \\
\pi_B^{\oplus}(i, k) - \pi_A^{\ominus}(i, k) &= \sum_{j=0}^{N/2-i-1} \frac{N!}{(i-k)!k!j!(i+j+1)!(N-2i-2j-1)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^j \sigma_B^j \sigma_{\phi}^{N-2i-2j-1} \cdot (\sigma_B^{i+1} - \sigma_A^{i+1}), \\
\pi_0^{\oplus}(i, k) - \pi_0^{\ominus}(i, k) &= \sum_{j=0}^{N/2-i} \frac{N!}{(i-k)!k!j!(i+j)!(N-2i-2j)!} \sigma_1^{i-k} \sigma_2^k \sigma_A^j \sigma_B^j \sigma_{\phi}^{N-2i-2j} \cdot (\sigma_B^i - \sigma_A^i).
\end{aligned}$$

Assume that $\sigma_A < \sigma_B$. It follows that $Er_A - Er_B < 0$ and thus the citizen should vote for candidate A which contradicts our assumption that $\sigma_A < \sigma_B$. For $\sigma_A > \sigma_B$ there is a corresponding contradiction. Thus, the strategies are consistent only for $\sigma_A = \sigma_B$.

The difference in expected loss between voting for A and abstaining is

$$\begin{aligned}
Er_A - Er_\phi &= \sum_{i=1}^{N/2} \sum_{k=0}^i \left\{ \int_0^\infty \left[\pi_A^\oplus(i, k) \cdot r^*(s, x(i, k)) + \pi_B^\oplus(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] + \pi_0^\oplus(i, k) \cdot r^*(s, x(i, k)) \right] f(s, x(i, k)) ds \right. \\
&+ \int_{-\infty}^0 \left[\pi_A^\ominus(i, k) \cdot \bar{r}(s, x(i, k)) + \pi_B^\ominus(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] + \pi_0^\ominus(i, k) \cdot \bar{r}(s, x(i, k)) \right] f(s, x(i, k)) ds \\
&- \int_0^\infty \left[\pi_A^\oplus(i, k) \cdot r^*(s, x(i, k)) + \pi_B^\oplus(i, k) \cdot \bar{r}(s, x(i, k)) + \pi_0^\oplus(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] \right] f(s, x(i, k)) ds \\
&\left. - \int_{-\infty}^0 \left[\pi_A^\ominus(i, k) \cdot \bar{r}(s, x(i, k)) + \pi_B^\ominus(i, k) \cdot r^*(s, x(i, k)) + \pi_0^\ominus(i, k) \cdot \left[\frac{1}{2}r^*(s, x(i, k)) + \frac{1}{2}\bar{r}(s, x(i, k)) \right] \right] f(s, x(i, k)) ds \right\} \\
&= \sum_{i=1}^{N/2} \sum_{k=0}^i \frac{1}{2} \left[\pi_B^\oplus(i, k) - \pi_B^\ominus(i, k) + \pi_0^\oplus(i, k) - \pi_0^\ominus(i, k) \right] \cdot \underbrace{\left(r^*(x(i, k)) - \bar{r}(x(i, k)) \right)}_{<0}.
\end{aligned}$$

With $\sigma_B = \sigma_A = \sigma$ it follows that $\pi_0^\oplus(i, k) - \pi_0^\ominus(i, k) = 0$ and

$$\begin{aligned}
\pi_B^\oplus(i, k) - \pi_B^\ominus(i, k) &= \sum_{j=0}^{N/2-i-1} \frac{N!}{(i-k)!k!j!(i+j)!(N-2i-2j-1)!} \sigma_1^{i-k} \sigma_2^k \sigma^{i+2j+1} \sigma_\phi^{N-2i-2j-1} \cdot \underbrace{\left(\frac{1}{i+j+1} - \frac{1}{j+1} \right)}_{<0} \\
&- \frac{N!}{(i-k)!k!(i-1)!(N-2i+1)!} \sigma_1^{i-k} \sigma_2^k \sigma^{i-1} \sigma_\phi^{N-2i+1} < 0.
\end{aligned}$$

Thus $Er_A - Er_B = 0$, $Er_A - Er_\phi > 0$ and $Er_B - Er_\phi > 0$. □

Proof of Lemma 3.

The value of the information service is equal to

$$\omega = E[\ell_{v^*}(\theta)] - E\{E[\ell_{v^*}(\theta|s)]\}$$

and we can rewrite this equation as

$$\begin{aligned}
\omega &= E\{E[\ell_{v^*}(\theta|s)]\} - E\{E[\ell_{v^*}(\theta|s)]\} \\
&= E\{E[\ell_{v^*}(\theta|s)] - E[\ell_{v^*}(\theta|s)]\} \\
&= E\{E[\ell_{v^*}(\theta|s) - \ell_{v^*}(\theta|s)]\}
\end{aligned}$$

where $E[\ell_{v^*}(\theta)] = E\{E[\ell_{v^*}(\theta|s)]\}$ in the first line can be verified by applying the law of iterated expectation. The next step is to inspect the term in curly brackets in the last line. Given the prior belief, the citizens are indifferent between the two candidates. We assume with no loss of generality that the citizens elect candidate A whenever they

are indifferent. The optimal posterior decision is

$$v^*(s) = \begin{cases} A & \text{for } E(\theta|s) \geq 0 \\ B & \text{for } E(\theta|s) < 0 \end{cases}$$

as described above. Substituting the optimal decisions into $\{E[\ell_{v^*}(\theta|s) - \ell_{v^*}(\theta|s)]\}$ we obtain

$$\begin{aligned} E[\ell_A(\theta|s) - \ell_{v^*}(\theta|s)] &= \begin{cases} E[\ell_A(\theta|s) - \ell_A(\theta|s)] & \text{for } E(\theta|s) \geq 0 \\ E[\ell_A(\theta|s) - \ell_B(\theta|s)] & \text{for } E(\theta|s) < 0 \end{cases} \\ &= \begin{cases} 0 & \text{for } E(\theta|s) \geq 0 \\ -E(\theta|s) & \text{for } E(\theta|s) < 0. \end{cases} \end{aligned}$$

We have already shown that $E(\theta|s) = \frac{\tau_\varepsilon x}{\tau_\theta + \tau_\varepsilon x} \cdot s$ and we can use this result to compute the posterior distribution of $E(\theta|s)$ as perceived from the prior perspective. From the prior perspective, $E(\theta|s)$ is the random variable

$$\frac{\tau_\varepsilon x}{\tau_\theta + \tau_\varepsilon x} \cdot (\theta + \varepsilon),$$

where $\theta \sim N(0, \tau_\theta)$ and $\varepsilon \sim N(0, \tau_\varepsilon x)$. Recalling the rules for sums of independently distributed normal random variables we have the result that $E(\theta|s)$ is normally distributed with mean $m = 0$ and precision $r = \frac{\tau_\theta(\tau_\theta + x\tau_\varepsilon)}{x\tau_\varepsilon}$. Let G denote the distribution function of $E(\theta|s)$. Then, the value of the information service is equal to

$$\begin{aligned} \omega &= \int_{-\infty}^0 -w \, dG(w) + \int_0^{\infty} 0 \, dG(w) \\ &= \int_{-\infty}^0 -w \, dG(w). \end{aligned}$$

This is the partial expectation of a normal random variable. Inserting m and r into the functional form defined by equation (4) yields

$$\omega = \int_{-\infty}^0 -wg(w)dw = r^{-\frac{1}{2}} \cdot \varphi(0).$$

Substituting for r we obtain

$$\omega(x) = \left(\frac{\tau_\theta(\tau_\theta + x\tau_\varepsilon)}{x\tau_\varepsilon} \right)^{-1/2} \cdot \varphi(0).$$

Concavity of $\omega(x)$.

The first and second derivatives are

$$\omega'(x) = \frac{1}{2} \left(\frac{\tau_\theta^2}{\tau_\varepsilon x} + \tau_\theta \right)^{-3/2} \varphi(0) \frac{\tau_\theta^2}{\tau_\varepsilon x^2} > 0$$

and

$$\begin{aligned} \omega''(x) &= \frac{1}{2} \varphi(0) \cdot \left[\frac{3}{2} r^{-\frac{5}{2}} \left(\frac{\tau_\theta^2}{\tau_\varepsilon x^2} \right)^2 - 2r^{-\frac{3}{2}} \frac{\tau_\theta^2}{\tau_\varepsilon x^3} \right] \\ &= \frac{1}{2} \varphi(0) \frac{\tau_\theta^2}{\tau_\varepsilon x^3} r^{-\frac{3}{2}} \left[\frac{-(\tau_\theta + 4x\tau_\varepsilon)}{2(\tau_\theta + \tau_\varepsilon x)} \right] < 0. \end{aligned}$$

Thus, $\omega(x)$ is a strictly concave function in x .

The slope of $\omega(x)$ at the origin.

We are interested in the limit value $\lim_{x \rightarrow 0} \omega'(x)$. To determine the limit value it is sufficient to inspect $\frac{1}{x^2} \cdot \left(\frac{\tau_\theta^2}{\tau_\varepsilon x} + \tau_\theta \right)^{-3/2}$. First we rearrange:

$$\left(\frac{\tau_\theta^2}{\tau_\varepsilon x} + \tau_\theta \right)^{-3/2} \cdot \frac{1}{x^2} = \left[\left(\frac{\tau_\theta^2}{\tau_\varepsilon x} + \tau_\theta \right) \cdot x^{4/3} \right]^{-3/2} = \left(\frac{\tau_\theta^2}{\tau_\varepsilon} x^{1/3} + \tau_\theta x^{4/3} \right)^{-3/2}$$

It is easy to see that

$$\lim_{x \rightarrow 0} \left(\frac{\tau_\theta^2}{\tau_\varepsilon} x^{1/3} + \tau_\theta x^{4/3} \right) = 0$$

and thus

$$\lim_{x \rightarrow 0} \left(\frac{\tau_\theta^2}{\tau_\varepsilon} x^{1/3} + \tau_\theta x^{4/3} \right)^{-3/2} = \infty$$

□

Proofs of the comparative statics analysis

The quality of the information service $x^* = x^*(\gamma, \tau_\theta, \tau_\varepsilon)$ is implicitly defined by the equation

$$\frac{1}{2} \left(\frac{\tau_\theta^2}{\tau_\varepsilon x} + \tau_\theta \right)^{-3/2} \varphi(0) \frac{\tau_\theta^2}{\tau_\varepsilon x^2} - \gamma = 0.$$

Though we cannot explicitly solve for x^* , comparative statics results can be derived by using the implicit-function rule. We define the function

$$z(x; \gamma, \tau_\theta, \tau_\varepsilon) := \frac{1}{2} \left(\frac{\tau_\theta^2}{\tau_\varepsilon x} + \tau_\theta \right)^{-3/2} \varphi(0) \frac{\tau_\theta^2}{\tau_\varepsilon x^2} - \gamma.$$

Given $z(x; \gamma, \tau_\theta, \tau_\varepsilon) = 0$ the implicit-function rule yields

$$\frac{\partial x}{\partial \alpha} = -\frac{z_\alpha}{z_x} \quad \text{for } \alpha \in \{\gamma, \tau_\theta, \tau_\varepsilon\}.$$

The denominator of $-\frac{z_\alpha}{z_x}$ is the second derivative $\omega''(x) = z_x$ and we have already shown above that $\omega''(x) < 0$. Hence, the sign of $\frac{\partial x}{\partial \alpha} = -\frac{z_\alpha}{z_x}$ is determined by the sign of z_α .

Effect of γ . Since $z_\gamma = -1$, we obtain the result that $\partial x^*/\partial \gamma = -z_\gamma/z_x < 0$. It follows for the individual total expected loss $\rho_t = \rho(x^*) + \gamma \cdot (x^*/(N+1))$ that

$$\begin{aligned} \frac{\partial \rho_t}{\partial \gamma} &= \frac{\partial \rho}{\partial x} \frac{\partial x}{\partial \gamma} + \frac{x}{(N+1)} + \frac{k}{(N+1)} \cdot \frac{\partial x}{\partial \gamma} \\ &= \underbrace{\left(\frac{\partial \rho}{\partial x} + \frac{\gamma}{(N+1)} \right)}_{<0} \underbrace{\frac{\partial x}{\partial \gamma}}_{<0} + \frac{x}{(N+1)} > 0 \end{aligned}$$

□

Effect of τ_ε . Computing z_{τ_ε} yields:

$$\begin{aligned} z_{\tau_\varepsilon} &= \frac{1}{2} \varphi(0) \cdot \frac{\tau_\theta^2}{x^2} \left[\frac{\partial r^{-\frac{3}{2}}}{\partial \tau_\varepsilon} \cdot \frac{1}{\tau_\varepsilon} - r^{-\frac{3}{2}} \cdot \frac{1}{\tau_\varepsilon^2} \right] \\ &= \frac{1}{2} \varphi(0) \cdot \frac{\tau_\theta^2}{x^2 \tau_\varepsilon^2} r^{-\frac{3}{2}} \left[\frac{\tau_\theta - 2\tau_\varepsilon x}{2(\tau_\theta + \tau_\varepsilon x)} \right], \end{aligned}$$

and it follows that

$$\frac{\partial x^*}{\partial \tau_\varepsilon} = \frac{x^*}{\tau_\varepsilon} \cdot \frac{\tau_\theta - 2\tau_\varepsilon x}{\tau_\theta + 4\tau_\varepsilon x} \begin{cases} > 0 & \text{for } \tau_\theta > 2\tau_\varepsilon x^* \\ = 0 & \text{for } \tau_\theta = 2\tau_\varepsilon x^* \\ < 0 & \text{for } \tau_\theta < 2\tau_\varepsilon x^*. \end{cases}$$

Hence, we find that

$$\frac{\partial(\tau_\varepsilon x^*)}{\partial \tau_\varepsilon} = -\frac{z_{\tau_\varepsilon}}{z_x} \cdot \tau_\varepsilon + x = 2x \cdot \frac{(\tau_\theta + \tau_\varepsilon x)}{\tau_\theta + 4\tau_\varepsilon x} > 0.$$

Further, we can show that the individual total expected loss decreases in τ_ε :

$$\begin{aligned} \frac{\partial \rho_t}{\partial \tau_\varepsilon} &= \frac{\partial \rho}{\partial x} \frac{\partial x}{\partial \tau_\varepsilon} + \frac{\gamma}{(N+1)} \cdot \frac{\partial x}{\partial \tau_\varepsilon} + \frac{\partial \rho}{\partial \tau_\varepsilon} \\ &= \frac{x}{\tau_\varepsilon(\tau_\theta + 4\tau_\varepsilon x)} \cdot \left[-\omega'(x) \cdot 2(\tau_\theta + \tau_\varepsilon x) + \frac{\gamma}{(N+1)} \cdot (\tau_\theta - 2\tau_\varepsilon x) \right] < 0. \end{aligned}$$

This follows because, in equilibrium, $\omega'(x) > \gamma/(N+1)$ and $|2(\tau_\theta + \tau_\varepsilon x)| > |\tau_\theta - 2\tau_\varepsilon x|$. □

Effect of τ_θ . Computing z_{τ_θ} yields:

$$\begin{aligned} z_{\tau_\theta} &= \frac{1}{2}\varphi(0) \frac{1}{x^2 \tau_\varepsilon} \left[\frac{\partial r^{-\frac{3}{2}}}{\partial \tau_\theta} \cdot \tau_\theta^2 + r^{-\frac{3}{2}} \cdot 2\tau_\theta \right] \\ &= \frac{1}{2}\varphi(0) \frac{1}{x^2 \tau_\varepsilon} \left[-\frac{3}{2} r^{-\frac{5}{2}} \left(\frac{2\tau_\theta + \tau_\varepsilon x}{\tau_\varepsilon x} \right) \cdot \tau_\theta^2 + r^{-\frac{3}{2}} \cdot 2\tau_\theta \right] \\ &= \underbrace{\frac{1}{2}\varphi(0) \frac{\tau_\theta}{x^2 \tau_\varepsilon} r^{-\frac{3}{2}}}_{>0} \left[\frac{-2\tau_\theta + \tau_\varepsilon x}{2(\tau_\theta + \tau_\varepsilon x)} \right], \end{aligned}$$

and it follows that

$$\frac{\partial x^*}{\partial \tau_\theta} = \frac{x^*}{\tau_\theta} \cdot \frac{-2\tau_\theta + \tau_\varepsilon x}{\tau_\theta + 4\tau_\varepsilon x} \begin{cases} > 0 & \text{for } \tau_\theta < \tau_\varepsilon x^*/2 \\ = 0 & \text{for } \tau_\theta = \tau_\varepsilon x^*/2 \\ < 0 & \text{for } \tau_\theta > \tau_\varepsilon x^*/2. \end{cases}$$

The effect of an increase of τ_θ on the total individual expected loss is:

$$\begin{aligned} \frac{\partial \rho_t}{\partial \tau_\theta} &= \frac{\partial E[\ell_{v^*}(\theta)]}{\partial \tau_\theta} - \left[\frac{\partial \omega}{\partial \tau_\theta} + \frac{\partial \omega}{\partial x} \frac{\partial x}{\partial \tau_\theta} \right] + \frac{\gamma}{(N+1)} \cdot \frac{\partial x}{\partial \tau_\theta} \\ &= -\frac{1}{2}\varphi(0)\tau_\theta^{-\frac{3}{2}} \cdot \left(1 - \frac{4(\tau_\theta + \tau_\varepsilon x)^{\frac{1}{2}}(\tau_\varepsilon x)^{\frac{1}{2}}}{(\tau_\theta + 4\tau_\varepsilon x)} \right) + \frac{\gamma}{(N+1)} \cdot \frac{\partial x}{\partial \tau_\theta} \end{aligned}$$

We define $\lambda := \left(1 - \frac{4(\tau_\theta + \tau_\varepsilon x)^{\frac{1}{2}} (\tau_\varepsilon x)^{\frac{1}{2}}}{(\tau_\theta + 4\tau_\varepsilon x)}\right)$ and a straightforward calculation shows that

$$\lambda \begin{cases} > 0 & \text{for } \tau_\theta > 8\tau_\varepsilon x^* \\ = 0 & \text{for } \tau_\theta = 8\tau_\varepsilon x^* \\ < 0 & \text{for } \tau_\theta < 8\tau_\varepsilon x^*. \end{cases}$$

Recalling our results concerning $\frac{\partial x}{\partial \tau_\theta}$ it follows that

$$\frac{\partial \rho_t}{\partial \tau_\theta} < 0 \text{ for } \tau_\theta / \tau_\varepsilon x^* > 8$$

and

$$\frac{\partial \rho_t}{\partial \tau_\theta} > 0 \text{ for } \tau_\theta / \tau_\varepsilon x^* < 1/2.$$

For the domain $1/2 < \tau_\theta / \tau_\varepsilon x^* < 8$, however, the sign of $\frac{\partial \rho_t}{\partial \tau_\theta}$ depends on the respective magnitude of both the change in the electoral expected loss $(-\frac{1}{2}\varphi(0)\tau_\theta^{-\frac{3}{2}} \cdot \lambda)$ and in the price of a subscription $\frac{\gamma}{(N+1)} \cdot \frac{\partial x}{\partial \tau_\theta}$. \square

References

- ALDASHEV, G. (2010), Political information acquisition for social exchange, *Quarterly Journal of Political Science* 5(1), 1-25.
- BATTAGLINI, M., MORTON, R.B. and PALFREY, T.R. (2010), The swing voter's curse in the laboratory, *Review of Economic Studies* 77(1), 61-89.
- BESLEY, T. and GHATAK, M. (2007), Retailing public goods: The economics of corporate social responsibility, *Journal of Public Economics* 91(9), 1645-1663.
- DEGROOT, M. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.
- DOWNES, A. (1957), *An Economic Theory of Democracy*, HarperCollins Publishers, New York.
- FEDDERSEN, T. and PESENDORFER, W. (1996), The swing voter's curse, *American Economic Review* 86, 408-424.
- FEDDERSEN, T. and SANDRONI, A. (2006), Ethical voters and costly information acquisition, *Quarterly Journal of Political Science* 1(3), 287-311.
- GERARDI, D. and YARIV, L. (2008), Information acquisition in committees, *Games and Economic Behavior* 62(2), 436-459.
- GERLING, K., GRÜNER, H.P., KIEL, A. and SCHULTE, E. (2005), Information acquisition and decision making in committees: A survey, *European Journal of Political Economy* 21(3), 563-597.
- GERSHKOV, A. and SZENTES, B. (2009), Optimal voting schemes with costly information acquisition, *Journal of Economic Theory* 144(1), 36-68.
- HORRIGAN, J., GARRETT, K. and RESNICK, P. (2004), *The Internet and Democratic Debate*, Pew Internet & American Life Project.
- LI, H. (2001), A theory of conservatism, *Journal of Political Economy* 109(3), 617-636.
- MARTINELLI, C. (2006), Would rational voters acquire costly information?, *Journal of Economic Theory* 129(1), 225-251.
- MARTINELLI, C. (2007), Rational ignorance and voting behavior, *International Journal of Game Theory* 35(3), 315-335.
- MARTINELLI, C. (2011), Ignorance and naivete in large elections, Working Papers 1107, Centro de Investigacion Economica, ITAM.

- MUKHOPADHAYA, K. (2003), Jury size and the free rider problem, *Journal of Law, Economics and Organization* 19(1), 24-44.
- OLIVEROS, S. (2011a), Abstention, ideology and information acquisition, UC Berkeley, mimeo.
- OLIVEROS, S. (2011b), Aggregation of endogenous information in large elections, UC Berkeley, mimeo.
- OLSON, M. (1965), *The Logic of Collective Action*, Harvard University Press, Cambridge, MA.
- PECORINO, P. (1999), The effect of group size on public good provision in a repeated game setting, *Journal of Public Economics* 72(1), 121-134.
- PERSICO, N. (2004), Committee design with endogenous information, *Review of Economic Studies* 71(1), 165-191.
- PRAT, A. and STRÖMBERG, D. (2011), The political economy of mass media, *CEPR Discussion Paper No. DP8246*.
- PRATT, J., RAIFFA, H. and SCHLAIFER, R. (1995), *Introduction to Statistical Decision Theory*, The MIT Press.
- STRÖMBERG, D. (2004), Mass media competition, political competition, and public policy, *Review of Economic Studies* 71(1), 265-284.